Use of redundant exclusion PCR to identify a novel Bacillus thuringiensis Cry8 toxin gene from pooled genomic DNA

Article (Accepted Version)

Use of Redundant Exclusion PCR to identify a novel *Bacillus thuringiensis* Cry8 toxin gene from pooled genomic DNA

Fengjiao Zhang[a, b], Changlong Shu[b], Neil Crickmore[c], Yanqiu Li[b], Fuping Song[b], Chunqin Liu[d], Zhibao Chen[a]#, Jie Zhang[b]#

[a] College of Life Science and Technology Heilongjiang Bayi Agricultural University, Daqing, 163000, P. R. China

[b] State Key Laboratory for Biology of Plant Diseases and Insect Pests, Institute of Plant Protection, Chinese Academy of Agricultural Sciences, Beijing, 100193, P. R. China

[c] School of Life Sciences, University of Sussex, Brighton, BN1 9QG, UK

[d] Cangzhou Academy of Agricultural and Forestry Sciences, Cangzhou 061001, P. R. China

Running title: Redundant Exclusion PCR to identify novel Bt gene

#Address correspondence to Zhibao Chen (chenzhibao_2000@sina.com, Tel: +86 459 6819293; Fax: +86 459 6819291) and Jie Zhang (jzhang@ippcaas.cn, Tel: +86 10 62896634; Fax: +86 10 62812642)

F. Z. and C. S. These authors contributed equally to this work.

Abstract

With the aim of optimizing the cloning of novel genes from a genomic pool containing many previously identified, homologous, genes we designed a redundant exclusion PCR technique. In RE-PCR a pair of generic amplification primers are combined with additional primers that are designed to specifically bind to redundant, unwanted genes that are a subset of those copied by the amplification primers. During RE-PCR the specific primer blocks amplification of the full length redundant gene. Using this method we managed to clone a number of *cry8* or *cry9* toxin genes from a pool of *Bacillus thuringiensis* genomic DNA while excluding amplicons for *cry9Da, cry9Ea* and *cry9Eb.* The method proved very efficient at increasing the number of rare genes in the resulting library. One such rare, and novel, *cry8*-like gene was expressed and the encoded toxin was shown to be toxic to *Anomola corpulenta.*

Importance

Protein toxins from the bacterium *Bacillus thuringiensis* are being increasingly used as biopesticides against a wide range of insect pests, yet the search for new or improved toxins is becoming more difficult as traditional methods for gene discovery routinely isolate previously identified clones. This paper describes an approach that we have developed to increase the success rate for novel toxin gene identification through reducing or eliminating the cloning of previously characterized genes.

Introduction

As a result of the proteinaceous insecticidal toxins produced by *Bacillus thuringiensis* (Bt) this bacterium has become a commercially successful biopesticide [1]. Products based on Bt include formulations of the bacterium itself or the toxin expressed in an alternative host – in particular genetically modified crops [2]. Despite the increasing use of these products there remains a need to discover new toxins with desirable properties - such properties include an increased activity against a given target, activity against a new target pest or the ability to control a pest that has developed resistance to an existing toxin. A number of different approaches can be used to identify novel toxins, the traditional one being to screen strains for a desired activity and then isolate the active ingredient. In recent times molecular approaches have been increasingly used including genome sequencing [3] and PCR techniques. The latter rely on there being conserved regions present in toxin gene families as well as the more variable regions that give toxins their individual characteristics [4]. Improved PCR procedures have allowed the successful cloning of Bt toxin genes from complex DNA mixtures prepared from pooled samples [5, 6]. A problem with this sort of approach however is the high ratio of known or undesired toxin genes in libraries made from these pooled samples which has made the discovery of new genes increasingly difficult.

B. thuringiensis Cry8 and Cry9 proteins have significantly different insecticidal spectra despite phylogenetic analyses indicating that they share high sequence similarity in domains I and II [7, 8]. Cry8 proteins are toxic to Coleopteran insects while Cry9s have high activity to Lepidopteran insects [9-12], both are valuable toxins for insect pest management. This paper describes a procedure developed to analyse *cry8* and *cry9* genes in a DNA pool prepared from

2000 Bt strains and used to efficiently clone novel isolates.

Materials and methods

Bacterial strains, plasmids, and growth conditions

Bt stains were isolated from soil samples in China as described previously [6]. *E. coli* DH5α was used for standard transformations, whilst *E.coli* SCS110 (RpsL [strr], thr, leu, endA, thi-1, lacY, galK, galT, ara, tonA, tsx, dam, dcm, upE44, D [lac-proAB] [F0 traD36 proAB lacIqZDM15]) was used to produce non-methylated plasmid DNA for transformation of HD73$^-$, a crystal-negative mutant strain of Bt.

The cloning vector pEB was constructed by inserting the pETblue-2 expression region into the *Sty*I-*Bgl*II sites of pET-21b. Plasmid pSTK containing the *cry3Aa* promoter and STAB-SD sequence was constructed by Wang et al. [13] and used to express *cry* genes in HD73$^-$.

E.coli was incubated at 37°C in LB medium (1% NaCl, 1% tryptone, and 0.5% yeast extract), and Bt strains were grown at 30°C on LB agar (1% NaCl, 1% tryptone, 0.5% yeast extract and 1.3%-1.5% agar). To select antibiotic-resistant *E.coli* and Bt strains, ampicillin and kanamycin were added to the culture medium at a final concentration of 100 μg/ml and 50 μg/ml respectively. All of the cultures were incubated in a rotary shaker at 220 rpm.

Identification of *cry8* and *cry9* genes from pooled genomic DNA

Pooled genomic DNA was prepared from 2000 Bt strains as described by Li et al [6] and used as the template for PCR. A pair of universal primers cry_F and cry_R (Table 1) were designed based on an alignment of *cry8* and *cry9* holotype genes and were used to amplify the

toxin coding region of these genes. A 50 µl PCR mixture contained 100 ng template DNA, 25 µl 2×PrimeSTAR master mix (TaKaRa, Dalian, China), and 0.2 μmol^{-1} of each primer. The reaction consisted of an initial denaturation step of 94°C for 5 min, 30 cycles of 94°C for 1min; 54°C for 1min; 72°C for 2 min 20 s, and final extension step of 72°C for 10 min. The resulting PCR product was purified using a DNA gel extraction kit (AXYGEN, Hangzhou, China) and ligated into the *Ecl*136II site of the pEB vector.

The *cry9* and *cry8* genes in the pooled DNA were classified by PCR-RFLP analysis. Genes were amplified from library clones using cry_F and cry_R, digested with *Hinf* I, and then profiled by 2% agarose gel electrophoresis. Clones with different restriction length fragment polymorphism (RFLP) profiles were selected for sequencing.

Redundant Exclusion PCR

Primers RE9Da_F and RE9Ea/b_F (Table 2) were designed to specifically hybridize to *cry9Da*, and to *cry9Ea/cry9Eb* respectively. Test reactions were performed in 20 µL total volume containing 10 ng of *cry9Ea* and *cry9Eb* or *cry9Da*-encoding plasmid DNA, 0.2 $\mu molL^{-1}$ primer RE9Da_F or RE9Ea/b_F, 0.2 µmol L^{-1} of primer cry_F and/or cry_R, and 10 µl 2×PrimeSTAR master mix. PCR consisted of an initial denaturation step of 94°C for 5 min, 30 cycles of 94°C 1min; 54°C 1min; 72°C, 2 min 20 s, and a final extension step of 72°C for10 min.

Expression of the *cry8*-like gene

To express the truncated *cry8*-like gene in Bt, a seamless assembly cloning method was used to fuse the truncated gene to DNA encoding the Cry8Ea C-terminal coding region. The primers designed to amplify these two sections, with appropriate overlaps, are listed in Table 3. A 10 µL

reaction mix containing 20 ng pSTK plasmid (linearized with *Bam*HI+*Sal*I), 30 ng each of the PCR products and 5 μL 2×Assembly Master Mix (Seamless Assembly Cloning kit, CloneSmarter, USA）was incubated at 50°C for 10 min. After transformation of *E. coli* DH5α the resulting hybrid (hy*cry8*) was sequenced using an automated DNA sequencer (ABI-3730XL, USA). The recombinant plasmid, isolated from *E.coli*, was used to transform SCS110 prior to introduction into Bt strain HD73^{-} by electroporation. A single transformant was selected from LB plates containing kanamycin (50 g/ml) and incubated until sporulation at 30°C. The spore-crystal mixture was washed and re-suspended in sterile distilled water, and the suspension was examined by microscopy and SDS-PAGE analysis according to Shu et al [14]. For proteolytic activation the toxin crystals were solubilized in 50mM Na_2CO_3 pH10 and then treated with chymotrypsin (10:1 w/w) in PBS and incubated at 37°C for 2 h.

Insect bioassay

The insecticidal activity of spore-crystal mix from the recombinant Bt strain was tested against larvae of 15-day-old *Anomala corpulenta*, 5-day-old *Holotrichia parallela*, and 5-day-old *Holotrichia oblita*. The bioassay diet for these scarab larvae was prepared as described by Yu et al [15]. For initial screening, we used a concentration of 1.0×10^8 colony forming units (CFU) g^{-1}soil. Further assays were performed with samples showing toxicity to any of the pests in order to determine LC_{50} values. Bioassays were repeated at least twice, and LC_{50} values were calculated using probit analysis.

Results

Identification of *cry8* and *cry9* genes from pooled genomic DNA

A library of 2.1 kb PCR products encoding the active portion of *cry8/9* toxin genes (Fig.1A lane1), produced from the pooled genomic DNA using primers cry_F and cry_R, was created and subjected to PCR-RFLP analysis. Two hundred clones were tested and four distinct profile types were detected (Fig.1B lane1-4). Representatives of these were sequenced and this indicated that the four profiles belonged to *cry9Da* (100% identity to *cry9Da4*, accession number: GQ249297.1), *cry9Ea* (99% identity to *cry9Ea9*, accession number: JN651495.1), *cry9Eb* (99% identity to *cry9Eb2*, accession number: GQ249298.1) and a new *cry8*-like (86% identity to *cry8Ab1*, accession number: EU044830.1). Figure 2 (solid bars) shows the relative frequency of these four profiles, 70% of the clones matched the *cry9Ea* profile while 5% matched *cry8*-like.

Redundant Exclusion primer PCR

To exclude known genes from the library we relied on the fact that the high fidelity polymerase that we used in our PCR reactions (PrimeSTAR GXL) lacks a 5'-3' exonuclease activity and so polymerization can be halted by the presence of a bound oligonucleotide. We designed primers that would specifically, and tightly, bind to particular *cry8/9* genes and thus prevent amplification of the full PCR product by the primers. The primers RE9Da_F and RE9Ea/b_F (Tables 2 and 4) were designed to hybridize to variable regions in domain II of *cry9Da* and *cry9Ea/9Eb* respectively. A number of PCRs were run to test these primers. When the two amplification primers and the two redundant exclusion primers were all included in a multiplex PCR (Fig1A lane 2) two distinct bands are seen, the 2.1kb band represents the full

length products amplified by cry_F and cry_R while the approximately 750 bp band represents the fragment of the gene amplified by RE9Da_F/RE9Ea/b_F and cry_R. The source of this 750 bp band was confirmed in separate reactions involving just an RE primer and cry_R (Fig 1A lanes 3 and 4). To confirm that the RE primers could block amplification of the corresponding gene pairs of reactions were performed using specific genes as templates. Figure 1A lanes 5-8 show that amplification of the full length gene is completely (*cry9Da* lane 8) or mostly (*cry9Ea*/*b* lane 6) inhibited by the appropriate RE primer.

A new library was created from the 2.1kb PCR product produced using cry_F and cry_R from pooled DNA but this time including the RE primers in the amplification reaction. A total of 200 clones were analyzed by PCR-RFLP and this time five different profiles were obtained (Fig.1B lanes 3-7). Two of these profiles (*cry9Ea* and *cry8*-like) were the same as previously identified while the other three corresponded to *cry8Fa* (99% identity to *cry8Fa2*, accession number HQ174208.1), *cry8Ab*-like (99% identity to *cry8Ab1*-like, accession number: JF521572.1) and *cry8Ea* (99% identity to *cry8Ea1*, accession number: AY329081.1). Although the *cry9Ea* profile was detected despite the presence of the corresponding RE primer its frequency had dropped dramatically (Figure 2 open bars). The *cry9Da* and *cry9Eb* profiles had successfully been excluded. In contrast the frequency of the *cry8*-like profile rose significantly.

Expression of the novel *cry8*-like gene

The cry_F and cry_R primers were designed to amplify the active toxin portion of the *cry8/9* genes but not the region encoding the C-terminal crystallization domain [4]. To express a complete toxin we added this C-terminal region from a homologous gene, *cry8Ea*. Figure 3 shows the sequence of the Cry8-like toxin and the resulting hybrid (hyCry8). The hybrid gene was

sub-cloned into pSTK and introduced into HD73$^-$ for expression. The hyCry8 protein expressed well in this host (Fig.4C lane 2) and accumulated as spherical crystals (Fig.4B).When the protoxin was treated with chymotrypsin (Fig.4C lane 3), a 60 kDa protein was obtained as expected [16].

Insect bioassay

To evaluate the toxicity of the hybrid protein, spore/crystal mix of the recombinant Bt strain was tested for insecticidal activity on larvae of 15-day stage *A.corpulenta*, 5-day stage *H.parallela*, and 5-day stage *H.oblita.* At a concentration of 1.0×10^8 CFU g^{-1} soil the parent strain HD73$^-$ had no insecticidal activity against any of the three bioassayed insects. The strain expressing the hybrid was shown to be toxic to *A.corpulenta* larvae (Fig.5) but not to *H.parallela* or *H.oblita*.

Discussion

In recent years many research projects have focused on cloning desired genes from complex DNA samples (metagenomic or pooled DNA) [5, 6, 17-19]. In such samples the distribution of gene homologues can be unequal and it can be a difficult task to isolate rare forms. Here we have demonstrated the use of redundant exclusion PCR to remove unwanted homologues from a gene pool and thus increase the frequency of rare forms in an amplicon library.

Bt is an insect pathogen which during sporulation produces insecticidal proteins that accumulate in the mother cell and form intracellular parasporal crystals [1, 20]. These crystals contain toxins, encoded by *cry* genes, which are noted for their specific insecticidal activity. Cry proteins have been expressed in crops to control agricultural pests [21, 22]. Due to their commercial value much research has focused on toxin gene discovery and here we have demonstrated the use of an RE-PCR based method that can identify new and rare *cry* gene

183 homologues from a genomic DNA pool and identified a novel toxin with activity against an
184 economically important pest. The method relies on us being able to identify unique regions in
185 redundant genes and would therefore exclude any otherwise novel genes that happened to share
186 that region. Nonetheless we have demonstrated that the principle works well and alternative RE
187 primers can always be designed in an attempt to overcome this potential limitation. The fact that
188 some *cry9Ea* profiles could still be detected may represent the fact that this process is not 100%
189 efficient and so there can be some background, which will be particularly noticeable for high
190 abundance genes. Alternatively polymorphisms in the RE primer region could allow amplification
191 of these variants.

192 The toxin that we identified by this method is closely related to other Cry8 toxins,
193 comparison of the active toxin region with all toxins in the nomenclature reveals that the closest
194 match is to Cry8Ab1 with 79% identity. This level of identity is sufficient to consider the toxin
195 novel - studies have shown that just a few amino acid differences in a toxin can not only determine
196 whether or not it is toxic, but also to which insect it has toxicity [23]. This conclusion is supported
197 by our bioassay data: the variant that we have isolated shows activity to *A. corpulenta* but not to *H.*
198 *oblita* or *H. parallela*. This contrasts with Cry8N which showed activity against the latter insect
199 but not the former two [24], with Cry8Ga which has activity against the two *Holotrichia* species
200 but not *A. corpulenta* [12], Cry8Ab with activity against the two Holotrichia species [25] and with
201 Cry8Fa and another Cry8-like protein that have activity against none or all of these three species
202 respectively [10, 11]. Thus on both sequence and insect specificity levels this newly described
203 toxin has novel attributes which have potential not only for the control of these insects but also in
204 the understanding of specificity determination. It should be noted that our bioassays were

conducted on a hybrid toxin in which the C-terminal region was provided by another toxin. While the C-terminal region of the larger Cry toxins is known to be important for toxin expression/packaging [26] and may influence toxicity [27] there is no convincing evidence that it influences specificity and thus we believe that it is reasonable to attribute the observed activity to the newly cloned portion of the toxin.

In conclusion, the RE-PCR based approach developed in this paper was able to effectively exclude undesired homologue genes but clone low frequency genes from a complex DNA sample. It is particularly applicable to families of homologous genes where there are known areas with variation that can be used for the design of the RE primers, the method should be suitable for other samples such as metagenomic DNA.

Ethical Statement

The authors declare that they have no competing interests; the manuscript does not contain experiments using mammals and does not contain studies on humans.

References

1. Bravo A, Likitvivatanavong S, Gill SS, Soberon M: *Bacillus thuringiensis*: A story of a successful bioinsecticide. *Insect biochemistry and molecular biology* 2011, 41(7):423-431.
2. Sanahuja G, Banakar R, Twyman RM, Capell T, Christou P: *Bacillus thuringiensis*: a century of research, development and commercial applications. *Plant biotechnology journal* 2011, 9(3):283-300.
3. Ye W, Zhu L, Liu Y, Crickmore N, Peng D, Ruan L, Sun M: Mining new crystal protein genes from *Bacillus thuringiensis* on the basis of mixed plasmid-enriched genome sequencing and a computational pipeline. *Applied and environmental microbiology* 2012, 78(14):4795-4801.
4. Schnepf E, Crickmore N, Rie JV, Lereclus D, Baum J, Feitelson J, Zeigler DR, Dean DH: *Bacillus thuringiensis* and its pesticidal crystal proteins. *Microbiology & Molecular Biology Reviews* 1998, 62(3):775-806.
5. Shu C, Zhang J, Chen G, Liang G, He K, Crickmore N, Huang D, Zhang J, Song F: Use of a pooled clone method to isolate a novel *Bacillus thuringiensis* Cry2A toxin with activity

against *Ostrinia furnacalis*. *Journal of invertebrate pathology* 2013, 114(1):31-33.

6. Li Y, Shu C, Zhang X, Crickmore N, Liang G, Jiang X, Liu R, Song F, Zhang J: Mining rare and ubiquitous toxin genes from a large collection of *Bacillus thuringiensis* strains. *Journal of invertebrate pathology* 2014, 122:6-9.
7. Bravo A: Phylogenetic relationships of *Bacillus thuringiensis* delta-endotoxin family proteins and their functional domains. *Journal of bacteriology* 1997, 179(9):2793-2801.
8. de Maagd RA, Bravo A, Crickmore N: How *Bacillus thuringiensis* has evolved specific toxins to colonize the insect world. *Trends in genetics : TIG* 2001, 17(4):193-199.
9. Shu C, Su H, Zhang J, He K, Huang D, Song F: Characterization of *cry9Da4*, *cry9Eb2*, and *cry9Ee1* genes from *Bacillus thuringiensis* strain T03B001. *Applied microbiology and biotechnology* 2013, 97(22):9705-9713.
10. Bi Y, Zhang Y, Shu C, Crickmore N, Wang Q, Du L, Song F, Zhang J: Genomic sequencing identifies novel *Bacillus thuringiensis* Vip1/Vip2 binary and Cry8 toxins that have high toxicity to Scarabaeoidea larvae. *Applied microbiology and biotechnology* 2015, 99(2):753-760.
11. Shu C, Yu H, Wang R, Fen S, Su X, Huang D, Zhang J, Song F: Characterization of two novel *cry8* genes from *Bacillus thuringiensis* strain BT185. *Current microbiology* 2009, 58(4):389-392.
12. Shu C, Yan G, Wang R, Jie Z, Feng S, Huang D, Song F: Characterization of a novel *cry8* gene specific to Melolonthidae pests: *Holotrichia oblita* and *Holotrichia parallela*. *Applied Microbiology & Biotechnology* 2009, 84(84):701-707.
13. Wang G, Zhang J, Song F, Wu J, Feng S, Huang D: Engineered *Bacillus thuringiensis* GO33A with broad insecticidal activity against lepidopteran and coleopteran pests. *Applied microbiology and biotechnology* 2006, 72(5):924-930.
14. Shu C, Liu R, Wang R, Zhang J, Feng S, Huang D, Song F: Improving toxicity of *Bacillus thuringiensis* strain contains the *cry8Ca* gene specific to *Anomala corpulenta* larvae. *Current microbiology* 2007, 55(6):492-496.
15. Yu H, Zhang J, Huang D, Gao JG, Song F: Characterization of *Bacillus thuringiensis* Strain Bt185 Toxic to the Asian Cockchafer: *Holotrichia parallela*. *Current microbiology* 2006, 53(1):13-17.
16. Bulla LA, Jr., Bechtel DB, Kramer KJ, Shethna YI, Aronson AI, Fitz-James PC: Ultrastructure, physiology, and biochemistry of *Bacillus thuringiensis*. *Critical reviews in microbiology* 1980, 8(2):147-204.
17. Owen JG, Zachary CP, Smith AG, Ternei MA, Calle PY, Reddy BVB, Daniel M, Brady SF: Multiplexed metagenome mining using short DNA sequence tags facilitates targeted discovery of epoxyketone proteasome inhibitors. *Proceedings of the National Academy of Sciences* 2015, 112(14):4221-4226.
18. Annette K, Gagunashvili AN, Gulder TAM, Morinaka BI, Cristina D, Markus G, Miao VPW, Jörn P, Ólafur S A: Metagenomic natural product discovery in lichen provides evidence for a family of biosynthetic pathways in diverse symbioses. *Proceedings of the National Academy of Sciences of the United States of America* 2013, 110(33):E3129-E3137.
19. Gong JS, Lu ZM, Li H, Zhou ZM, Shi JS, Xu ZH: Metagenomic technology and genome mining: emerging areas for exploring novel nitrilases. *Applied Microbiology & Biotechnology* 2013, 97(15):6603-6611.

20. Raymond B, Johnston PR, Christina NL, Didier L, Neil C: *Bacillus thuringiensis*: an impotent pathogen? *Trends in microbiology* 2010, 18(5):189–194.
21. Kumar S, Chandra A, Pandey KC: *Bacillus thuringiensis* (Bt) transgenic crop: an environment friendly insect-pest management strategy. *Journal of environmental biology / Academy of Environmental Biology, India* 2008, 29(5):641-653.
22. Bates SL, Zhao JZ, Roush RT, Shelton AM: Insect resistance management in GM crops: past, present and future. *Nature biotechnology* 2005, 23(1):57-62.
23. Abdullah MA, Alzate O, Mohammad M, McNall RJ, Adang MJ, Dean DH: Introduction of Culex toxicity into *Bacillus thuringiensis* Cry4Ba by protein engineering. *Applied and environmental microbiology* 2003, 69(9):5343-5353.
24. Li H, Liu R, Shu C, Zhang Q, Zhao S, Shao G, Zhang X, Gao J: Characterization of one novel *cry8* gene from *Bacillus thuringiensis* strain Q52-7. *World journal of microbiology & biotechnology* 2014, 30(12):3075-3080.
25. Yuan Z, Zheng G, Tan J, Li C, Cheng L: Cloning and characterization of a novel *cry8Ab1* gene from *Bacillus thuringiensis* strain B-JJX with specific toxicity to scarabaeid (Coleoptera: Scarabaeidae) larvae. *Microbiological research* 2013, 168(8):512-517.
26. Evdokimov AG, Moshiri F, Sturman EJ, Rydel TJ, Zheng M, Seale JW, Franklin S: Structure of the full-length insecticidal protein Cry1Ac reveals intriguing details of toxin packaging into in vivo formed crystals. *Protein science : a publication of the Protein Society* 2014, 23(11):1491-1497.
27. Gomez I, Sanchez J, Munoz-Garay C, Matus V, Gill SS, Soberon M, Bravo A: *Bacillus thuringiensis* Cry1A toxins are versatile proteins with multiple modes of action: two distinct pre-pores are involved in toxicity. *The Biochemical journal* 2014, 459(2):383-396.

301 Table 1 Homologues of *cry9* and *cry8* genes and the design of universal amplification primers.

Primers and Genes	Sequence	Location	Genbank Acc number
cry_F	ATGAATCGAAATAATCAAAATGAATAT		
cry9Bb1	ATGAATCGAAATAATCAAAATGAATAT	1~27	X75019.1
cry9Ca1	ATGAATCGAAATAATCAAAATGAATAT	1~27	Z37527.1
cry9Da1	ATGAATCGAAATAATCAAAATGAATAT	1~27	D85560.1
cry9Db1	ATGAATCGAAATcATCAAAATGAATAT	1~27	AY971349.1
cry9Dc1	ATGAATCGAAATAATCAAAATGAATAT	1~27	KC156683.1
cry9Ea1	ATGAATCGAAATAATCcAAATGAATAT	1~27	AB011496.1
cry9Eb1	ATGAAcCGAAATAATCAAAATGAtTAT	1~27	AX189653.1
cry9Ec1	ATGAATCGAAATAATCAAAATGAATAT	1~27	AF093107.2
cry9Ed1	ATGAATCGAAATAATCAAAATGAATAT	1~27	AY973867.1
cry9Ee1	ATGAATCGAAATAATCAAAATGAATAT	1~27	GQ249296.1
cry9Fa1	ATGAcTaGAAATAgACAAgATGAATAT	1~27	KC156692.1
cry8Aa1	ATGAgTCcAAATAATCAAAATGAATAT	1~27	U04364.1
cry8Ab1	ATGAgTCcAAATAATCAAAATGAATAT	1~27	EU044830.1
cry8Ac1	ATGAgTCcAAATAATCAAAATGAATAT	1~27	KC156662.1
cry8Ad1	ATGAgTCcAAATAATCAAAATGAATAT	1~27	KC156684.1
cry8Ba1	ATGAgTCcAAATAATCAAAATGAATAT	1~27	U04365.1
cry8Bb1	ATGAgTCcAAATAATCAAAATGAATAT	1~27	AX543924.1
cry8Bc1	ATGAgTCcAAATAATCAAAATGAATAT	1~27	AX543926.1
cry8Ca1	ATGAgTCgAAATAATCAAAATGAgTAT	1~27	U04366.1
cry8Db1	ATGAgTCcAAATAATCAAAATGAATAT	1~27	AB303980.1
cry8Ea1	ATGAgTCcAAATAATCAAAATGAATAT	1~27	AY329081.1
cry8Fa1	ATGAgTCcAAATAATCAAAATGAATAT	1~27	AY551093.1
cry8Ga1	ATGAgTCcgAATAATCAgAAcGAATAT	1~27	AY590188.1
cry8Ha1	ATGAaTCcgAATAATCAgAATGAATAT	1~27	AY897354.2
cry8Ia1	ATGAgTCcgAATAATCAgAATGAgTtT	1~27	EU381044.1

cry8Ib1	ATGAgcCcAAATAATCAAAATGAgTtT	1~27	GU325772.1
cry8Ja1	ATGAgTCcgAATAATCAgAATGAgTAT	1~27	EU625348.1
cry8Ka1	ATGAgTCcAAATAATCtAAATGAATAT	1~27	FJ422558.1
cry8Kb2	ATGAgTCcAAATAATCAAAATGAATAT	1~27	FJ770570.1
cry8Na1	ATGAgTCcgAATAATCAAAAcGAATAT	1~27	HM640939.1
cry8Pa1	ATGAgTCcAAATAATCAAAATGAATAT	1~27	HQ388415.1
cry8Qa1	ATGAgTCcAAATAATCAAAATGAATAT	1~27	HQ441166.1
cry8Ta1	ATGAgTCaAAATAATCAAAATGAATAT	1~27	KC156673.1
cry_R	GATAAGCAYGACACTAAATTTGC		
cry9Bb1	GCAAATTTAGTGTtATGCTTATC	2110~2132	X75019.1
cry9Ca1	GCAAATTTAGTGTCATGCTTATC	2092~2114	Z37527.1
cry9Da1	GCAAATTTAGTGTCATGCTTATC	2125~2147	D85560.1
cry9Db1	GCAAATTTAGTGTCATGCTTATC	2128~2150	AY971349.1
cry9Dc1	GCAAATTTAGTGTCGTGCTTATC	2125~2147	KC156683.1
cry9Ea1	GCAAATTTAGTGTCGTGCTTATC	2071~2093	AB011496.1
cry9Eb1	GCAAATTTAGTGTCATGCTTATC	2074~2096	AX189653.1
cry9Ec1	GCAAATTTAGTGTCATGCTTATC	2083~2105	AF093107.2
cry9Ed1	GCAAATTTAGTGTCGTGCTTATC	2083~2105	AY973867.1
cry9Ee1	GCAAATTTAGTGTCGTGCTTATC	2089~2111	GQ249296.1
cry9Fa1	GCAAATTTAGTGTCATGCTTAaC	2089~2111	KC156692.1
cry8Aa1	GCAAAcTTAGTGGaATGCcTATC	2104~2126	U04364.1
cry8Ab1	GCAAAcTTAGTGGaATGCcTATC	2122~2144	EU044830.1
cry8Ac1	GCAAAcTTAGTGGaATGCcTATC	2146~2168	KC156662.1
cry8Ad1	GCcAAcTTAGTGGaATGCcTATC	2104~2126	KC156684.1
cry8Ba1	GCcAAcTTAGTGGaATGCcTATC	2092~2114	U04365.1
cry8Bb1	GCAAAcTTAGTGGaATGCcTATC	2107~2129	AX543924.1
cry8Bc1	GCAAAcTTAGTGGaATGCcTATC	2119~2141	AX543926.1
cry8Ca1	GCAAAcTTAaTagaATGCcTATC	2095~2117	U04366.1
cry8Db1	GCAAAcTTAGTagaATGCcTATC	2137~2159	AB303980.1

cry8Ea1	GCAAAcTTAGTGgaATGCcTATC	2074~2096	AY329081.1
cry8Fa1	GCAAAcTTAGTGgaATGCcTATC	2104~2126	AY551093.1
cry8Ga1	GCAAAcTTAGTagaATGCcTATC	2087~2109	AY590188.1
cry8Ha1	GCtAATTTAGTagaATGCcTATC	2089-2111	AY897354.2
cry8Ia1	GCAAATTTAaTtgaATGCgTATC	2113~2135	EU381044.1
cry8Ib1	GCAAATTTAaTtgaATGCgTATC	2156~2138	GU325772.1
cry8Ja1	GCAAAcTTAaTagaATGCcTATC	2101~2123	EU625348.1
cry8Ka1	GCAAAcTTAGTcgaATGCcTATC	2086~2108	FJ422558.1
cry8Kb2	GCcAAcTTAGTGgaATGCcTATC	2092~2114	FJ770570.1
cry8Na1	GCAAAcTTAGTagaATGCcTATC	2104~2126	HM640939.1
cry8Pa1	GCAAAcTTAGTcgaATGCcTATC	2095~2117	HQ388415.1
cry8Qa1	GCAAAcTTAGTcgaATGCcTATC	2119~2141	HQ441166.1
cry8Ta1	GCAAATTTAaTtgaATGCgTATC	2149~2171	KC156673.1

302

303

304 Table 2 Homologues of *cry9* and *cry8* genes and the design of redundant exclusion primers.

Primers and Genes	Sequence	Location	Genbank Acc number
RE9Da_F	AGGATATACACAGCAAGGTATACC AGC		
cry9Bb1	--AACGAA------CTTTGTTA------	1346~1359	X75019.1
cry9Ca1	--GACTTC----TCCTGCTAATGG-AGG	1323~1343	Z37527.1
cry9Da1	AGGATATACACAGCAAGGTATACC-AGC	1347~1373	D85560.1
cry9Db1	-GGACAAA---ATAACGTTCTTCC-ACA	1366~1388	AY971349.1
cry9Dc1	-GGACAAA---ATAACGTTTTGCC-ACC	1366~1388	KC156683.1
cry9Ea1	--CACCAA----TGCAGCTAATAC-GTG	1335~1355	AB011496.1
cry9Eb1	--CACCTC----AGCCCCTAATAC-GTG	1335~1355	AX189653.1
cry9Ec1	--CATTTC----TCCTGCTAATGC-AGG	1332~1352	AF093107.2
cry9Ed1	--CATTTC----TCCTGCTAATGC-AGG	1332~1352	AY973867.1
cry9Ee1	--GACTCA----ACCTTCTACTGG-AGG	1332~1352	GQ249296.1
cry9Fa1	--GTCAAC-----CCACCTAATTC-TGG	1345~1364	KC156692.1
cry8Aa1	TATATTCAAAAACACATACAGCTCTCCA	1346~1373	U04364.1
cry8Ab1	TTTATTCTAAAACACATACAACTGGAGA	1349~1376	EU044830.1
cry8Ac1	TTTATTCTAAAACATATACAACTCCAAA	1349~1376	KC156662.1
cry8Ad1	TTTATTCAAAAACACATACAACTCCATA	1343~1370	KC156684.1
cry8Ba1	AACGTATAAACCAGCTTCCAAAGATATT	1347~1374	U04365.1
cry8Bb1	AAAGTATAATCCAGTTTCCAAAGATATT	1347~1374	AX543924.1
cry8Bc1	AAAGTATAATCCGGTTTCCAAAGATATT	1347~1374	AX543926.1
cry8Ca1	CTTAT-TCGAAGCCAAAACAATTC-GCG	1322~1347	U04366.1
cry8Db1	CGTACTCAAAACCACATACAACTATGCA	1352~1379	AB303980.1
cry8Ea1	CCTATAAT-----CCTG-GATCTGAAGG	1328~1379	AY329081.1
cry8Fa1	CTCATTTTTTCTGATAG-TACGGGAGGG	1330~1357	AY551093.1
cry8Ga1	GGTATCAAAAAGAATCTA-ATGTC-CCA	1322~1347	AY590188.1
cry8Ha1	TGGATACGATATAGCGTTTAGCGAAA--	1332~1357	AY897354.2

cry8Ia1	TAATTATGAACCTCCAGGCATATCCA-A	1329~1355	EU381044.1
cry8Ib1	TGAATATGATCTTCAACTTTTGTCTA-A	1332~1358	GU325772.1
cry8Ja1	TTTACCTATAATCCTGGATCTGAA-GGT	1324~1350	EU625348.1
cry8Ka1	ATGAAAAAT-----TATCGAACTT----	1328~1346	FJ422558.1
cry8Kb2	ATGAAAAAT-----CATCGAACTT----	1328~1346	FJ770570.1
cry8Na1	TCTATCTTGTGGGGTG-----GTG-C--	1354~1373	HM640939.1
cry8Pa1	AGTGTATAAGCCGGTTTCCAAAGATATT	1341~1368	HQ388415.1
cry8Qa1	CTCACTTTCTCTGATAGTACGGGCGGAA	1327~1354	HQ441166.1
cry8Ta1	CGTATAGTAAAACCCATACAGCTATACA	1346~1373	KC156673.1
RE9Ea/b_F	GAAAT CACCAA TGCAGCTAATAC GT		
cry9Bb1	AACTC--------AACGAA------CTTTGTTA----	1341~1359	X75019.1
cry9Ca1	GGTAC--------GACTTC----TCCTGCTAATGG-AG	1318~1342	Z37527.1
cry9Da1	GGGATT---TCAGGATATACACAGCAAGGTATACC-AG	1339~1372	D85560.1
cry9Db1	CGTATG---TC-GGACAAA---ATAACGTTCTTCC-AC	1358~1379	AY971349.1
cry9Dc1	CGTATG---TC-GGACAAA---ATAACGTTTTGCC-AC	1358~1379	KC156683.1
cry9Ea1	GAAAT--------CACCAA----TGCAGCTAATAC-GT	1330~1354	AB011496.1
cry9Eb1	GAAAT--------CACCTC----AGCCCCTAATAC-GT	1330~1354	AX189653.1
cry9Ec1	GGTAC--------CATTTC----TCCTGCTAATGC-AG	1327~1351	AF093107.2
cry9Ed1	GGTAC--------CATTTC----TCCTGCTAATGC-AG	1327~1351	AY973867.1
cry9Ee1	GGCAC--------GACTCA----ACCTTCTACTGG-AG	1327~1351	GQ249296.1
cry9Fa1	AGTGTT-------GTCAAC-----CCACCTAATTC-TG	1363~1387	KC156692.1
cry8Aa1	AACAGCGTATTTATATTCAAAAACACATACAGCTCTCC	1335~1372	U04364.1
cry8Ab1	ATCATCTCATCTTTATTCTAAAACACATACAACTGGAG	1338~1375	EU044830.1
cry8Ac1	ATCAACTCAACTTTATTCTAAAACATATACAACTCCAA	1338~1375	KC156662.1
cry8Ad1	ATCATATTATTTTTATTCAAAAACACATACAACTCCAT	1332~1369	KC156684.1
cry8Ba1	AAGACG---TTAACGTATAAACCAGCTTCCAAAGATAT	1339~1373	U04365.1
cry8Bb1	AAGACG---TTAAAGTATAATCCAGTTTCCAAAGATAT	1339~1373	AX543924.1
cry8Bc1	AAGACG---TTAAAGTATAATCCGGTTTCCAAAGATAT	1339~1373	AX543926.1
cry8Ca1	AAAAA-----ACTTAT-TCGAAGCCAAAACAATTC-GC	1316~1347	U04366.1

cry8Db1	GGTTTT----ACGTACTCAAAACCACATACAACTATGC	1345~1378	AB303980.1
cry8Ea1	AACATT---TACCTATAAT-----CCTG-GATCTGAAG	1320~1348	AY329081.1
cry8Fa1	TGCACACA-CCCTCATTTTTTCTGATAG-TACGGGAGG	1320~1355	AY551093.1
cry8Ga1	GACCTT---TAGGTATCAAAAAGAATCTA-ATGTC-CC	1314~1347	AY590188.1
cry8Ha1	GGGATTGATGTTGGATACGATATAGCGTTTAGCGAAA-	1321~1357	AY897354.2
cry8Ia1	GGGAA----TTTAATTATGAACCTCCAGGCATATCCA-	1322~1344	EU381044.1
cry8Ib1	TAGAT----TATGAATATGATCTTCAACTTTTGTCTA-	1325~1357	GU325772.1
cry8Ja1	AACAAC----ATTTACCTATAATCCTGGATCTGAA-GG	1317~1349	EU625348.1
cry8Ka1	GAGTTACATGTATGAAAAAT-----TATCGAACTT---	1317~1346	FJ422558.1
cry8Kb2	GAGTTACAGGTATGAAAAAT-----CATCGAACTT---	1317~1346	FJ770570.1
cry8Na1	GGCCAC---GTTCTATCTTGTGGGGTG-----GTG-C-	1346~1373	HM640939.1
cry8Pa1	AAG------TTAGTGTATAAGCCGGTTTCCAAAGATAT	1336~1367	HQ388415.1
cry8Qa1	TGCCCCCA-ATCTCACTTTCTCTGATAGTACGGGCGGA	1317~1353	HQ441166.1
cry8Ta1	TGGCTCCCTTACGTATAGTAAAACCCATACAGCTATAC	1335~1372	KC156673.1

305

306 Table 3 Amplification Primers

Primer	Sequence (5'to3')
hycry8_F	TGGTGGACAGCAAATGGGTCGGGATCCGATGAATCGAAATAATC
hycry8_R	CATTTGGATACAAATCATCCGATAAGCATGACACTAAATTTGCCGC
hycry8Ea_F	GGATGATTTGTATCCAAATG
hycry8Ea_R	CTCGAGTGCGGCCGCAAGCTTGTCGACTTACTCTACGTCAACAATC

307

308

309

Applied and Environmental Microbiology

310 Table 4 Primers and Annealing Temperatures

Primers	A-base content	C-base content	G-base content	T-base content	length（bps）	TM（℃）
REDa_F	40.7%	22.2%	22.2%	14.8%	27	62.0
RE9Ea/b_F	27.6%	17.2%	27.6%	27.6%	29	68.0
cry_F	55.6%	7.4%	11.1%	25.9%	27	57.7
cry_R	39.1%	21.7%	17.4%	21.7%	23	55.2

311

312

313

314 Fig.1 PCR product restriction fragment length polymorphism profiles of cloned genes.

315 A:Lane1: PCR product from pooled genomic DNA and cry_F/cry_R primers; Lane2: PCR product

316 from pooled genomic DNA, cry_F/cry_R and REcry9Da_F/ REcry9Ea/b_F primers; Lane3: PCR

317 product from pooled genomic DNA and REcry9Da_F/cry_R primers; Lane4: PCR product from

318 pooled genomic DNA and REcry9Ea/b_F and cry_R primers; Lane5: PCR product from *cry9Ea*

319 and *cry9Eb* template and cry_F/cry_R primers； Lane6: PCR product from *cry9Ea* and *cry9Eb*

320 template, cry_F/cry_R and REcry9Ea/b_F primers； Lane7: PCR product from *cry9Da* template,

321 cry_F/cry_R； Lane8: PCR product from *cry9Da* template cry_F/cry_R and REcry9Da_F primers;

322 M: molecular weight marker (DL5000).

323 B: Lanes1-7:RFLP profile of cloned *cry9Eb, cry9Da, cry9Ea*, *cry8*-like, *cry8Fa*, *cry8Ab*-like,

324 *cry8Ea* genes. M: molecular weight marker (DL2000).

325

326 Fig.2 The proportion of clones in the genomic libraries. The solid bars represent clones isolated

327 from the normal library while the open bars represent those from the redundant exclusion library.

328

329 Fig.3 Amino acid alignment of Cry8Ab1, Cry8-like, hyCry8 and Cry8Ea.

330

331 Fig.4 SEM and SDS-PAGE analysis of spore/crystal mixtures. A: Bt strain HD73-; B: Bt strain

332 expressing hyCry8; C: SDS-PAGE. M: protein marker; lane1: HD73^{-}; lane2: hyCry8; lane3:

333 chymotrypsin-treated hybrid.

334

335 Fig.5 Insecticidal activity of a spore/crystal mixture of the hybrid expressing strain